CLAIMS

What is claimed is:

1. A method for concealing errors in an audio signal, comprising:

digitally encoding the audio signal into a plurality of audio data packets representative of the audio signal;

determining a perceptually tolerable distortion limit for said audio packets; and

altering a value of at least one said audio packet by an amount within said perceptually tolerable distortion limit utilizing information representative of a different said audio data packet.

2. A method in accordance with Claim 1 wherein a plurality of said audio packets are altered by an amount within said perceptually tolerable distortion, each alteration utilizing information representative of a different said audio packet than the audio packet being altered.

3. A method in accordance with Claim 2 wherein said alteration comprises fragile watermarking.

4. A method in accordance with Claim 3 wherein said alteration comprises least bit modulation (LBM).

5. A method in accordance with Claim 2 wherein said encoded audio data packets comprise modulated discrete cosine transform (MDCT) coefficients.

6. A method in accordance with Claim 5 wherein said altering a value of at least one said audio packet comprises modifying quantized indices of said encoded audio data packets.

7. A method in accordance with Claim 5 wherein said alteration comprises modulo watermarking.

8. A method in accordance with Claim 5 wherein said coefficients include coefficients corresponding to a plurality of bands within a time frame and said encoded audio data packets comprise a plurality of time frames, and wherein, for a band *i* and a time frame *n*, a coefficient is written $b[n,k]$, where $k \in \mathrm{K_i}$ and $\mathrm{K_i}$ is an index set of band i, and coefficient $b[n,k]$ includes two least significant bits having an integer value of 0, 1, 2, or 3 written $d[n,i]$,

and further wherein said altering at least one audio data packet comprises:

determining indices $c[n,i] = \mathrm{argmin}_{c\in\{0,1,2,3\}} \sum_{k\in \mathrm{K_i}} (b[n,k] - \hat{b}_c[n,k])^2$, wherein

$\hat{b}_0[n,k] = 0$, $\hat{b}_1[n,k] = b[n-1,k]$, $\hat{b}_2[n,k] = b[n+1,k]$, and

$\hat{b}_3[n,k] = \frac{1}{2}(b[n-1,k] + b[n+1,k])$; and

setting $d[n,i]=\begin{cases}0, & \text{if } c[n-1,i]\in\{0,1\}\wedge c[n+1,i]\in\{0,2\},\\ 1, & \text{if } c[n-1,i]\in\{2,3\}\wedge c[n+1,i]\in\{0,2\},\\ 2 & \text{if } c[n-1,i]\in\{0,1\}\wedge c[n+1,i]\in\{1,3\},\\ 3 & \text{if } c[n-1,i]\in\{2,3\}\wedge c[n+1,i]\in\{1,3\}.\end{cases}$

9. A method in accordance with Claim 5 wherein said coefficients include coefficients that are quantization indices corresponding to a plurality of bands within a time frame and said encoded audio data packets comprise a plurality of time frames, and wherein, for a band *i* and a time frame *n*, a quantization index is written $q[n,k]$, where $k \in \mathsf{K}_i$ and K_i is an index set of band i, and coefficient $b[n,k]$ includes least significant bits written $d[n,i]$, and further wherein said determining a perceptually tolerable distortion limit comprises determining a number K of different embeddable values, and $l=\sum_{k\in\mathsf{K}_i} q[n,k]-d[n,i] \bmod K$;

and further comprising:

selecting a lower limit $I_{\min}$ in accordance with a minimum quantization index for which distortion can be tolerated and selecting an upper limit $I_{\max}$ to prevent quantization indices from being outside a bound after modification;

and further wherein said altering at least one audio data packet comprises:

searching for l or $k-l$ of said quantization indices having the largest magnitude from all said quantization indices that lie within a range $[I_{\min,}I_{\max}]$, depending upon whether $0\le l<K/2$ or $K>l>K/2$, respectively;

when fewer than the searched for said quantization indices are found, leaving said found quantization indices unchanged, otherwise subtracting

or adding 1 from each said found quantization index depending upon whether $0 \le l < K/2$ or $K > l > K/2$.

10. A method in accordance with Claim 5 further comprising preselecting a frame offset k; and further wherein said altering at least one audio data packet comprises embedding a 1 or a 0 in a least significant bit of a coefficient in a frame $n+k$ of a band j, depending upon whether $\sum_i (X_i^j(n) - X_i^j(n-1))^2 > \sum_i (X_i^j(n))^2$, where $X_i^j(n)$ represents an *i*th coefficient of a subband *j* in a frame *n* produced by said digital encoding of the audio data.

11. A method for concealing errors in an audio signal, comprising:

decoding a digitally encoded audio signal, wherein said digitally encoded audio signal includes a plurality of audio data packets representative of the audio signal, and said plurality of audio data packets includes a plurality of altered audio data packets; wherein each said altered audio data packet comprises an alteration indicative of information representative of a different said audio data packet, and each said alteration is limited to a predetermined perceptually tolerable distortion limit;

determining that at least one said audio data packet is missing or unavailable from the digitally encoded audio signal;

extracting information representative of said missing or unavailable audio data packet from an alteration of at least one different, available audio data packet; and

utilizing said extracted information to estimate said missing or unavailable audio data packet.

12. A method in accordance with Claim 11 wherein more than one audio data packet is missing or unavailable, and said extracting and utilizing steps are iterated for each missing data packet.

13. A method in accordance with Claim 12 wherein said extracted information comprises a fragile watermark.

14. A method in accordance with Claim 13 wherein said extracted information comprises least bit modulation (LBM).

15. A method in accordance with Claim 12 wherein said altered audio data packets comprise altered modulated discrete cosine transform (MDCT) coefficients.

16. A method in accordance with Claim 15 wherein said coefficients include coefficients corresponding to a plurality of bands within a time frame and said encoded audio data packets comprise a plurality of time frames, and wherein, for a band *i* and a time frame *n*, said altered audio data packets comprise a coefficient written $b[n,k]$, where $k \in \mathrm{K_i}$ and $\mathrm{K_i}$ is an index set of band i, wherein coefficient $b[n,k]$ includes two least significant bits having an

integer value of 0, 1, 2, or 3 written $d[n,i]$, and further wherein $d[n,i]$ is altered so that

$$d[n,i]=\begin{cases}0, & \text{if } c[n-1,i]\in\{0,1\}\wedge c[n+1,i]\in\{0,2\},\\ 1, & \text{if } c[n-1,i]\in\{2,3\}\wedge c[n+1,i]\in\{0,2\},\\ 2 & \text{if } c[n-1,i]\in\{0,1\}\wedge c[n+1,i]\in\{1,3\},\\ 3 & \text{if } c[n-1,i]\in\{2,3\}\wedge c[n+1,i]\in\{1,3\},\end{cases}$$

where $c[n,i]=\operatorname{argmin}_{c\in\{0,1,2,3\}}\sum_{k\in K_i}(b[n,k]-\hat{b}_c[n,k])^2$, and $\hat{b}_0[n,k]=0$,

$\hat{b}_1[n,k]=b[n-1,k]$, $\hat{b}_2[n,k]=b[n+1,k]$, and $\hat{b}_3[n,k]=\frac{1}{2}(b[n-1,k]+b[n+1,k])$;

and further wherein:

said extracting information representative of said missing or unavailable audio data packet comprises extracting $d[n,i]$ for a plurality of time frames n; and

said utilizing said extracted information to estimate said missing or unavailable audio data packet comprises utilizing bits of said extracted $d[n,i]$ to determine whether to estimate a missing or unavailable coefficient utilizing a neighboring time frame.

17. A method in accordance with Claim 15 wherein said coefficients include coefficients that are quantization indices corresponding to a plurality of bands within a time frame and said encoded audio data packets comprise a plurality of time frames, and wherein, for a band *i* and a time frame *n*, a quantization index is written $q[n,k]$, where $k \in K_i$ and K_i is an index set of band i, and coefficient $b[n,k]$ includes least significant bits written $d[n,i]$, and further

wherein said predetermined perceptually tolerable distortion limit includes K different embeddable values, and $l = \sum_{k \in K_1} q[n,k] - d[n,i] \bmod K$;

and further wherein said extracting information representative of said missing or unavailable audio data packet comprises decoding $\hat{d}[n,i]$ as $\sum_{k \in K_1} q[n,k] \bmod K$.

18. A method in accordance with Claim 15 wherein, for a preselected frame offset k; said altered data packets comprise an embedded 1 or a 0 in a least significant bit $B(j)$ of a coefficient in a frame $n+k$ of a band j, depending upon whether $\sum_i (X_i^J(n) - X_i^J(n-1))^2 > \sum_i (X_i^J(n))^2$, where $X_i^J(n)$ represents an *i*th coefficient of a subband *j* in a frame *n* produced by said digital encoding of the audio data, wherein said least significant bits $B(j)$ are embedded for each *j* from 1 to J, wherein *j* is the band in which the bit is embedded, and *J* is the number of bands;

and for a lost frame n, said extracting information representative of said missing or unavailable audio data packet comprises extracting, from a frame $n+k$, embedded bits $B(j)$ for $j = 1, J$; and said utilizing said extracted information comprises estimating coefficient value $X_i^J(n)$ as either $X_i^J(n-1)$ or 0, depending upon the extracted embedded bits.

19. An apparatus for concealing errors in an audio signal, said apparatus configured to:

digitally encode the audio signal into a plurality of audio data packets representative of the audio signal; and

utilizing a determined perceptually tolerable distortion limit for said audio packets, alter a value of at least one said audio packet by an amount within said perceptually tolerable distortion limit utilizing information representative of a different said audio data packet.

20. An apparatus in accordance with Claim 19 configured to alter a plurality of said audio packets by an amount within said perceptually tolerable distortion, and

for each said alteration, utilize information representative of a different said audio packet than the audio packet being altered.

21. An apparatus in accordance with Claim 20 wherein said alteration comprises a fragile watermarking.

22. An apparatus in accordance with Claim 21 wherein said alteration comprises least bit modulation (LBM).

23. An apparatus in accordance with Claim 20 configured to encode said audio data packets as data including modulated discrete cosine transform (MDCT) coefficients.

24. An apparatus in accordance with Claim 23 wherein said coefficients include coefficients corresponding to a plurality of bands within a time frame and said encoded audio data packets comprise a plurality of time frames, and wherein, for a band *i* and a time frame *n*, a coefficient is written $b[n,k]$, where $k \in \mathrm{K_i}$ and $\mathrm{K_i}$ is an index set of band i, and coefficient $b[n,k]$ includes two least significant bits having an integer value of 0, 1, 2, or 3 written $d[n,i]$,

and further wherein to alter at least one audio data packet, said apparatus is configured to:

determine indices $c[n,i] = \mathrm{argmin}_{c\in\{0,1,2,3\}} \sum_{k\in \mathrm{K_i}} (b[n,k] - \hat{b}_c[n,k])^2$, wherein

$\hat{b}_0[n,k] = 0$, $\hat{b}_1[n,k] = b[n-1,k]$, $\hat{b}_2[n,k] = b[n+1,k]$, and

$\hat{b}_3[n,k] = \frac{1}{2}(b[n-1,k] + b[n+1,k])$; and

set $d[n,i] = \begin{cases} 0, & \text{if } c[n-1,i] \in \{0,1\} \wedge c[n+1,i] \in \{0,2\}, \\ 1, & \text{if } c[n-1,i] \in \{2,3\} \wedge c[n+1,i] \in \{0,2\}, \\ 2 & \text{if } c[n-1,i] \in \{0,1\} \wedge c[n+1,i] \in \{1,3\}, \\ 3 & \text{if } c[n-1,i] \in \{2,3\} \wedge c[n+1,i] \in \{1,3\}. \end{cases}$

25. An apparatus in accordance with Claim 23 wherein said coefficients include coefficients that are quantization indices corresponding to a plurality of bands within a time frame and said encoded audio data packets comprise a plurality of time frames, and wherein, for a band *i* and a time frame *n*, a quantization index is written $q[n,k]$, where $k \in \mathrm{K_i}$ and $\mathrm{K_i}$ is an index set of band i, and coefficient $b[n,k]$ includes least significant bits written $d[n,i]$, and further having a selected number K of different embeddable values, where

$l = \sum_{k \in K_i} q[n,k] - d[n,i] \bmod K$; a lower limit $I_{\min}$ in selected accordance with a minimum quantization index for which distortion can be tolerated; and an upper limit $I_{\max}$ to prevent quantization indices from being outside a bound after modification;

and further wherein to alter said at least one audio data packet, said apparatus is configured to:

search for l or $k-l$ of said quantization indices having the largest magnitude from all said quantization indices that lie within a range $[I_{\min,} I_{\max}]$, depending upon whether $0 \le l < K/2$ or $K > l > K/2$, respectively; and

when fewer than the searched for said quantization indices are found, leave said found quantization indices unchanged, otherwise subtract or add 1 from each said found quantization index depending upon whether $0 \le l < K/2$ or $K > l > K/2$.

26. An apparatus in accordance with Claim 23 and further wherein to alter at least one audio data packet, said apparatus is configured to embed a 1 or a 0 in a least significant bit of a coefficient in a frame $n+k$ of a band j, depending upon whether $\sum_i (X_i^j(n) - X_i^j(n-1))^2 > \sum_i (X_i^j(n))^2$, wherein $X_i^j(n)$ represents an *i*th coefficient of a subband *j* in a frame *n* produced by said digital encoding of the audio data; and further wherein k is a preselected frame offset.

27. An apparatus for concealing errors in an audio signal, said apparatus configured to:

decode a digitally encoded audio signal, wherein said digitally encoded audio signal includes a plurality of audio data packets representative of the audio signal, and said plurality of audio data packets includes a plurality of altered audio data packets; wherein each said altered audio data packet comprises an alteration indicative of information representative of a different said audio data packet, and each said alteration is limited to a predetermined perceptually tolerable distortion limit;

determine when at least one said audio data packet is missing or unavailable from the digitally encoded audio signal;

extract information representative of said missing or unavailable audio data packet from an alteration of at least one different, available audio data packet; and

utilize said extracted information to estimate said missing or unavailable audio data packet.

28. An apparatus in accordance with Claim 27 wherein more than one audio data packet is missing or unavailable, said apparatus configured to iterate said extracting and utilizing for each missing data packet.

29. An apparatus in accordance with Claim 28 configured to extract a fragile watermark.

30. An apparatus in accordance with Claim 29 configured to extract least bit modulation (LBM).

31. An apparatus in accordance with Claim 28 configured to decode altered audio data packets comprising altered modulated discrete cosine transform (MDCT) coefficients.

32. An apparatus in accordance with Claim 31 wherein said coefficients include coefficients corresponding to a plurality of bands within a time frame and said encoded audio data packets comprise a plurality of time frames, and wherein, for a band *i* and a time frame *n*, said altered audio data packets comprise a coefficient written $b[n,k]$, where $k \in \mathrm{K_i}$ and $\mathrm{K_i}$ is an index set of band i, wherein coefficient $b[n,k]$ includes two least significant bits having an integer value of 0, 1, 2, or 3 written $d[n,i]$, and further wherein $d[n,i]$ is altered so that

$$d[n,i] = \begin{cases} 0, & \text{if } c[n-1,i] \in \{0,1\} \wedge c[n+1,i] \in \{0,2\}, \\ 1, & \text{if } c[n-1,i] \in \{2,3\} \wedge c[n+1,i] \in \{0,2\}, \\ 2 & \text{if } c[n-1,i] \in \{0,1\} \wedge c[n+1,i] \in \{1,3\}, \\ 3 & \text{if } c[n-1,i] \in \{2,3\} \wedge c[n+1,i] \in \{1,3\}, \end{cases}$$

where $c[n,i] = \operatorname{argmin}_{c\in\{0,1,2,3\}} \sum_{k\in \mathrm{K_i}} (b[n,k] - \hat{b}_c[n,k])^2$, and $\hat{b}_0[n,k] = 0$,

$\hat{b}_1[n,k] = b[n-1,k]$, $\hat{b}_2[n,k] = b[n+1,k]$, and $\hat{b}_3[n,k] = \frac{1}{2}(b[n-1,k] + b[n+1,k])$;

and further wherein:

to extract information representative of said missing or unavailable audio data packet, said apparatus is configured to extract $d[n,i]$ for a plurality of time frames n; and

to utilize said extracted information to estimate said missing or unavailable audio data packet, said apparatus is configured to utilize bits of said extracted $d[n,i]$ to determine whether to estimate a missing or unavailable coefficient utilizing a neighboring time frame.

33. An apparatus in accordance with Claim 31 wherein said coefficients include coefficients that are quantization indices corresponding to a plurality of bands within a time frame and said encoded audio data packets comprise a plurality of time frames, and wherein, for a band *i* and a time frame *n*, a quantization index is written $q[n,k]$, where $k \in \mathrm{K_i}$ and $\mathrm{K_i}$ is an index set of band i, and coefficient $b[n,k]$ includes least significant bits written $d[n,i]$, and further wherein said predetermined perceptually tolerable distortion limit includes K different embeddable values, and $l = \sum_{k\in \mathrm{K_i}} q[n,k] - d[n,i] \bmod K$;

and further wherein to extract information representative of said missing or unavailable audio data packet, said apparatus is configured to decode $\hat{d}[n,i]$ as $\sum_{k\in \mathrm{K_i}} q[n,k] \bmod K$.

34. An apparatus in accordance with Claim 31 wherein, for a preselected frame offset k; said altered data packets comprise an embedded 1 or a 0 in a

least significant bit $B(j)$ of a coefficient in a frame $n+k$ of a band j, depending upon whether $\sum_i (X_i^j(n) - X_i^j(n-1))^2 > \sum_i (X_i^j(n))^2$, where $X_i^j(n)$ represents an *i*th coefficient of a subband *j* in a frame *n* produced by said digital encoding of the audio data, wherein said least significant bits $B(j)$ are embedded for each *j* from 1 to J, wherein *j* is the band in which the bit is embedded, and *J* is the number of bands;

and for a lost frame n, to extract information representative of said missing or unavailable audio data packet, said apparatus is configured to extract, from a frame $n+k$, embedded bits $B(j)$ for $j = 1, J$; and to utilize said extracted information, said apparatus is configured to estimate coefficient value $X_i^j(n)$ as either $X_i^j(n-1)$ or 0, depending upon the extracted embedded bits.

35. A machine readable medium having recorded thereon instructions configured to instruct a computer to:

digitally encode the audio signal into a plurality of audio data packets representative of the audio signal; and

utilizing a determined perceptually tolerable distortion limit for said audio packets, alter a value of at least one said audio packet by an amount within said perceptually tolerable distortion limit utilizing information representative of a different said audio data packet.

36. A machine readable medium having recorded thereon instructions configured to instruct a computer to:

decode a digitally encoded audio signal, wherein said digitally encoded audio signal includes a plurality of audio data packets representative of the audio signal, and said plurality of audio data packets includes a plurality of altered audio data packets; wherein each said altered audio data packet comprises an alteration indicative of information representative of a different said audio data packet, and each said alteration is limited to a predetermined perceptually tolerable distortion limit;

determine when at least one said audio data packet is missing or unavailable from the digitally encoded audio signal;

extract information representative of said missing or unavailable audio data packet from an alteration of at least one different, available audio data packet; and

utilize said extracted information to estimate said missing or unavailable audio data packet.